

Peptide Ion Fragmentation for Proteomics Applications

Tandem mass spectrometry is a central technology in the interdisciplinary, health-related field of proteomics. Great initial progress has been made by using data mining techniques to discern the most common patterns of peptide ion fragmentation. This knowledge has made it possible to infer sequence from the observed spectra. However, about half the information in the mass spectra is currently discarded as indecipherable. The goal of this project is to understand, at a predictive level, the chemical reactions responsible for peptide ion cleavage. This will improve the reliability and speed of proteomics measurements.

K.K. Irikura (Div. 838)

Current “informatics” approaches (statistical pattern recognition) have been effective for developing rules for interpreting the tandem mass spectra of peptides. The most important rules are pervasive and were discovered by using modest databases. Finding additional rules will make it possible to extract more sequence information from the same mass spectra. Unfortunately, these lower-frequency rules cannot be discovered empirically without enormous databases, which are not yet available. Thus, we have adopted a different strategy in this project.

Tandem mass spectrometry is essentially a pyrolytic process. A mass-selected peptide ion is subjected to multiple collisions with an unreactive gas. The collisions increase the ion’s internal energy until it finally dissociates. The observed mass spectrum is determined by the branching among the various unimolecular dissociation reactions.

In this project, NIST researchers studied individual peptide ions to understand their fragmentation patterns in quantitative, chemical detail. By understanding the underlying organic chemistry, we will be able to extract fragmentation rules directly, without massive databases.

Our technical approach involves integrated theory and experiment. The initial energy deposition process is modeled theoretically as a collisional energy transfer process. Reaction mechanisms, thermochemistry, and rate constants are obtained from *ab initio* electronic structure theory. Rate constants will also be measured experimentally using a novel instrument (still under construction) designed to attain complete ion thermalization, as required for com-

parison with statistical theories. Peptides designed specifically to reveal the underlying chemistry are synthesized and their tandem mass spectra are measured in the laboratory using standard instrumentation. Peptides are selected to answer questions raised by our theoretical calculations, and calculations are designed to answer questions raised by our experiments. After the theoretical thermal rate constants are validated experimentally, non-thermal rate constants will be computed by using the theoretical energy deposition functions. The resulting branching fractions will be compared directly with spectra collected under a variety of experimental conditions.

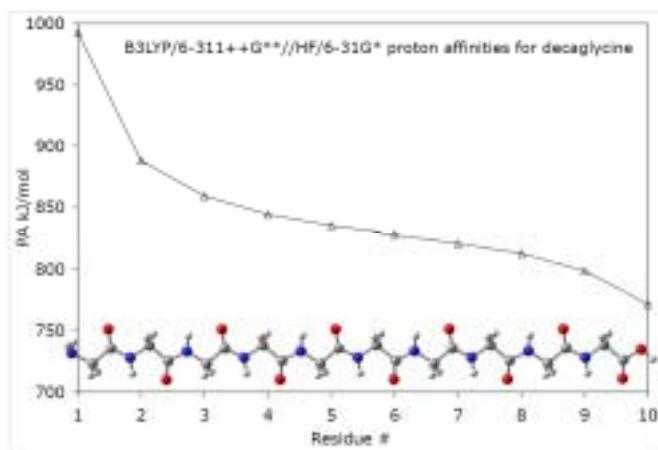
Synthesis and Measurements of Homologous Series of Peptides. A database of about 40,000 tandem mass spectra of yeast-derived peptides has been assembled so far. We search it continually for patterns of reactivity that have not been previously documented or explained. Suspected patterns are then investigated by designing and synthesizing a series of peptides that will best reveal the underlying chemistry. This is important because the peptides selected by *S. cerevisiae* are not expected to be particularly well suited for revealing reactivity patterns.

One such pattern is related to loss of neutral H₂O or NH₃ from peptides with an N-terminal glutamine (Gln) residue. This was initially recognized among the statistics for the yeast peptides. A series of Gln-initiated peptides was prepared with systematic variations of basicity and chain length. Tandem mass spectra of the synthetic peptides confirmed the suspected pattern and placed it on a quantitative basis. Another neutral loss detected in library search is the loss of 62 Da from peptide ions containing two threonines at the amino end. By examining the tandem mass spectra of a series of synthetic peptides as a function of collision energy it became clear that the loss takes place in two steps and involves an initial loss of water from the penultimate threonine followed by loss of acetaldehyde from the terminal one. The effect of aspartic acid on peptide fragmentation is being studied.

Reaction Mechanisms, Thermochemistry, and Rate Constants. Quantum chemistry calculations are traditionally applied to much smaller molecules. Our initial work has been to develop software tools to address the particular problems that arise for polypeptides. A procedure has been developed that combines both force-field and *ab initio* methods to find the most important conformational isomers for arbitrary gas-phase peptide ions. Applications to small models for the Gln chemistry described above gave only partial agreement with the experimental results. The next step is to run additional experiments to determine

the most efficient, chemically appropriate molecules for theoretical modeling.

The most popular chemical model in this field involves the "mobile proton hypothesis." Surprisingly little is known about the proton mobilization process. We are using both 1D lattice Coulomb models and *ab initio* polyglycine models to investigate the peculiar fragmentation of triply-charged peptide ions. Electro-static confinement of mobile protons is predicted to be an important effect; this will be tested experimentally. An unexpected trend in backbone proton affinities, shown in the figure below, was discovered for linear polyglycine. This biases fragmentation toward the N-terminus and can mask Coulombic effects beyond about 15 residues. Similar results have been obtained recently in aqueous solution; intramolecular electrostatic effects on half-wave potentials are damped beyond distances of about 1 nm.



There is little consensus about the chemical reaction mechanisms involved in peptide ion fragmentation. In addition to investigating mechanisms that have already been suggested (such as the mobile proton mechanism), we are using the NIST technique of isopotential searching (IPS) to discover novel mechanisms. Since peptide ions are large molecules, the existing IPS software requires weeks of computer time to yield results. To accelerate these calculations, the software is now being completely rewritten to take advantage of the large NIH cluster computer. We expect near-linear scaling, which should permit us to obtain results 10 to 50 times faster than now possible.

Experimentally, the construction and design of the thermalized peptide machine continue to proceed as planned. The rate constants for thermal (not collisional) peptide fragmentation will be measured. Their internal energy distribution will be known (Maxwell-Boltzmann), which is necessary to make corresponding theoretical predictions of unimolecular rate constants. Thus, this experiment makes it possible to test the quality of the *ab initio* derived rate constants.

Collisional Energy Transfer. Quasiclassical trajectory simulations are being performed to understand how translational energy is converted to vibrational and rotational energies. The results of trajectory calculations are used in Monte Carlo modeling of the resultant vibrational and rotational energy distributions in ions. Statistical rate theories are then used to evaluate the rates of ion fragmentation. Early results indicate that the modeling results are in semi-quantitative agreement with experiment. Energy distributions resulting from trajectory calculations are very sensitive to the details of the short-range repulsive part of the ion-collider potential energy surface. Surprisingly, the best agreement with experiment is obtained when hard-spheres collision model is used, with more realistic "softer" potentials resulting in lower ion dissociation rates.

Impact: The chemical reactions studied so far are fairly specific, so impact on proteomics is correspondingly minor. More general conclusions are anticipated as additional reactions are discovered. Insights about the fragmentation of triply-charged ions will be analytically useful.

Future Plans: All aspects of this project appear to be progressing well and will be continued: the size effect on activation energies will be investigated; atomistic molecular dynamics trajectories will be used to model collisional energy deposition; *ab initio* methods will be used to further characterize proton shuttling, backbone reactivity, and neutral losses; the isopotential-searching software will be parallelized; development will continue on the unique instrument for measuring thermal peptide fragmentation kinetics; more spectra will be added to the peptide mass spectral database.

Project Team: K.K. Irikura, J. K. Merle, S. E. Stein, P. Neta, V. D. Knyazev, and J. W. Hudgens (Div. 838)

Publications in preparation:

1. P. Neta, Q.-L. Pu, L. Kilpatrick, X. Yang, S. E. Stein *Dehydration vs. Deamination of N-Terminal Glutamine in Collision-Induced Dissociation of Protonated Peptides*; *J. Am. Soc. Mass Spectrom.*, in press.
2. P. Neta, Q.-L. Pu, X. Yang, S. E. Stein *Consecutive Neutral Losses of H₂O and C₂H₄O from N-Terminal Thr-Thr and Thr-Ser in Collision-Induced Dissociation of Protonated Peptides. Position Dependent Water Loss from Single Thr or Ser*; *Int. J. Mass Spectrom.*, submitted.
3. C. R. Kinsinger, K. K. Irikura *Efficient Conformational Searching of Protonated Peptides. How Good Are Force Fields for Gas-Phase Protonated Peptides?*; *J. Comput. Chem.*, submitted.
4. J. K. Merle, K. K. Irikura *Coulombic Confinement in the Mobile Proton Model*, Manuscript in preparation.